1. Show that, for any nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{C}^m$, the product $\mathbf{u}\,\mathbf{v}^H$ is a rank one matrix in $\mathbb{C}^{m\times m}$. Also show that, if $\mathbf{v}^H\,\mathbf{u} \neq -1$, then

$$\mathbf{I} - \frac{\mathbf{u}\,\mathbf{v}^H}{\mathbf{v}^H\mathbf{u}+1} = \left(\mathbf{I} + \mathbf{u}\,\mathbf{v}^H\right)^{-1}$$

(Note that $\left(\mathbf{I} + \mathbf{u}\,\mathbf{v}^H\right)$ is commonly called a *rank one perturbation of the identity.*)

solution:

Let $\mathbf{A} = \mathbf{u}\,\mathbf{v}^H$. Then, since $\mathbf{u} \neq \mathbf{0}$, there is a sequence of elementary row operations and a corresponding product of elementary matrices, denoted $\mathbf{E}$ such that

$$\mathbf{E}\,\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

But then

$$\mathbf{E}\,\mathbf{A} = \mathbf{E}\left(\mathbf{u}\mathbf{v}^H\right) = (\mathbf{E}\mathbf{u})\,\mathbf{v}^H = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathbf{v}^H = \begin{bmatrix} \mathbf{v}^H \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

This, however, implies that there is a sequence of elementary row operations that will reduce $\mathbf{u}\mathbf{v}^H$ to an echelon form with precisely one (nonzero) pivot row. Since the rank is, by definition, the number of pivot rows in the echelon form, then $\mathbf{u}\mathbf{v}^H$ is then of rank one.

If $\mathbf{v}^H\,\mathbf{u} \neq -1$, then direct computation shows

$$\left(\mathbf{I} - \frac{\mathbf{u}\,\mathbf{v}^H}{\mathbf{v}^H\mathbf{u}+1}\right)\,\left(\mathbf{I} + \mathbf{u}\,\mathbf{v}^H\right) = \mathbf{I} - \frac{\mathbf{u}\,\mathbf{v}^H}{\mathbf{v}^H\mathbf{u}+1} + \mathbf{u}\,\mathbf{v}^H - \frac{\mathbf{u}\,\mathbf{v}^H\mathbf{u}\,\mathbf{v}^H}{\mathbf{v}^H\mathbf{u}+1}$$

But note that $\mathbf{v}^H\,\mathbf{u}$ is a scalar, and can therefore be moved right or left in the last term. So the expression becomes

$$\begin{aligned}
\mathbf{I} - &\frac{1}{\mathbf{v}^H\mathbf{u}+1}\,\mathbf{u}\,\mathbf{v}^H + \mathbf{u}\,\mathbf{v}^H - \frac{\mathbf{v}^H\mathbf{u}}{\mathbf{v}^H\mathbf{u}+1}\,\mathbf{u}\,\mathbf{v}^H \\
&= \mathbf{I} + \left(-\frac{1}{\mathbf{v}^H\mathbf{u}+1} + 1 - \frac{\mathbf{v}^H\mathbf{u}}{\mathbf{v}^H\mathbf{u}+1}\right)\,\mathbf{u}\,\mathbf{v}^H \\
&= \mathbf{I} + (0)\mathbf{u}\,\mathbf{v}^H = \mathbf{I}
\end{aligned}$$

solution:

Therefore, since $\mathbf{u}\,\mathbf{v}^H$ is square, then by definition

$$\mathbf{I} - \frac{\mathbf{u}\,\mathbf{v}^H}{\mathbf{v}^H\mathbf{u} + 1} = \left(\mathbf{I} + \mathbf{u}\,\mathbf{v}^H\right)^{-1}$$

2. Consider the following matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix} \quad , \quad \mathbf{B} = \begin{bmatrix} 1 & 2 & -1 \\ -2 & 1 & 1 \\ 1 & -2 & 1 \end{bmatrix} \quad , \quad \mathbf{C} = \begin{bmatrix} \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{2} \\ \frac{3}{6} & \frac{1}{2} \\ \frac{5}{6} & -\frac{1}{2} \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \quad , \quad \mathbf{E} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{i}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{i}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{i}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{i}{2} \end{bmatrix}$$

where $i = \sqrt{-1}$. Which, if any, of these are unitary? Which, if any, are orthogonal? Of those that are neither, which are *easily* converted to unitary ones?

solution:

a. Consider $\mathbf{A}$. Note that $\mathbf{A}$ is real and symmetric. Therefore, by definition:

$$\mathbf{A}^H\,\mathbf{A} = \mathbf{A}^T\,\mathbf{A} = \mathbf{A}\ \mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$$

Direct computation shows that

$$\mathbf{A}^H\,\mathbf{A} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

Since this matrix is not the identity, then, by definition, $\mathbf{A}$ is neither unitary nor orthogonal. However, because the product is a diagonal matrix, then the columns of $\mathbf{A}$ are orthogonal, but not orthonormal vectors. Therefore, in order to create a matrix which is both unitary and orthogonal, we simply need to normalize each column of the original matrix. This produces:

$$\tilde{\mathbf{A}} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

which is easily shown to satisfy $\tilde{\mathbf{A}}^H\,\tilde{\mathbf{A}} = \tilde{\mathbf{A}}^T\,\tilde{\mathbf{A}} = \mathbf{I}$.

solution:

b. Consider **B**. Note that **B** is real, but not symmetric. Therefore, by definition:

$$\mathbf{B}^H\,\mathbf{B} = \mathbf{B}^T\,\mathbf{B} = \begin{bmatrix} 1 & -2 & 1 \\ 2 & 1 & -2 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ -2 & 1 & 1 \\ 1 & -2 & 1 \end{bmatrix} \quad ,$$

Direct computation the shows:

$$\mathbf{B}^H\,\mathbf{B} = \begin{bmatrix} 6 & -2 & -2 \\ -2 & 9 & -3 \\ -2 & -3 & 3 \end{bmatrix}$$

This result is not an indentity, and is not even diagonal. Therefore **B** is neither unitary nor orthogonal, and moreover there is no easy way (at least at this point) to create a unitary matrix from **B**.

solution:

c. Consider **C**. Note that **C** is real, but not square. Therefore, by definition,:

$$\mathbf{C}^H\,\mathbf{C} = \mathbf{C}^T\,\mathbf{C} = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{3}{6} & \frac{5}{6} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{2} \\ \frac{3}{6} & \frac{1}{2} \\ \frac{5}{6} & -\frac{1}{2} \end{bmatrix}$$

Direct computation the shows:

$$\mathbf{C}^H\,\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This result is an indentity. Therefore, the columns of **C** are orthonormal vectors. However, **C** is *not square*, Therefore, according to the definition in our text, **C** can be called neither unitary nor orthogonal. (Although the columns of **C** might be a good starting place from which to build a 4×4 unitary matrix!)

solution:

d. Note that $\mathbf{D}$ is complex. Therefore, by definition, $\mathbf{D}$ cannot be orthogonal, since this term applies only to real matrics. However:

$$\mathbf{D}^H\,\mathbf{D} = \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Since this matrix is not the identity, then, by definition, $\mathbf{D}$ is not unitary. However, because, as in part a., the product is a diagonal matrix, then the columns of $\mathbf{D}$ are orthogonal, but not orthonormal vectors, and, again we can create a matrix which is unitary by simply normalizing each column of the original matrix. This produces:

$$\tilde{\mathbf{D}} = \begin{bmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{i}{\sqrt{2}} \\ \dfrac{i}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{bmatrix}$$

which is easily shown to satisfy $\tilde{\mathbf{D}}^H\,\tilde{\mathbf{D}} = \mathbf{I}$.

solution:

e. Consider $\mathbf{E}$. Note that $\mathbf{E}$ is complex. Therefore, by definition, $\mathbf{E}$ cannot be orthogonal, since this term applies only to real matrics. However, we can show directly that:

$$\mathbf{E}^H\,\mathbf{E} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{i}{2} & -\frac{i}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{i}{2} & -\frac{i}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{i}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{i}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{i}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{i}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Since the product $\mathbf{E}^H\,\mathbf{E}$ produces the identity, then, by definition, $\mathbf{E}$ is unitary.

3. Consider the matrix:

$$\mathbf{C} = \begin{bmatrix} \frac{1}{6} & \frac{5}{6} \\ \frac{3}{6} & -\frac{1}{6} \\ \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & \frac{3}{6} \end{bmatrix}$$

Show that $\mathbf{C}^H\mathbf{C} = \mathbf{I}$, but $\mathbf{C}\mathbf{C}^H \neq \mathbf{I}$ and so $\mathbf{C}^H \neq \mathbf{C}^{-1}$. Explain why this is not a contradiction.

solution:

Direct computation shows

$$\mathbf{C}^H\mathbf{C} = \begin{bmatrix} \frac{1}{6} & \frac{3}{6} & \frac{5}{6} & \frac{1}{6} \\ \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{3}{6} \end{bmatrix} \begin{bmatrix} \frac{1}{6} & \frac{5}{6} \\ \frac{3}{6} & -\frac{1}{6} \\ \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & \frac{3}{6} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}$$

but

$$\mathbf{C}\mathbf{C}^H = \begin{bmatrix} \frac{1}{6} & \frac{5}{6} \\ \frac{3}{6} & -\frac{1}{6} \\ \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & \frac{3}{6} \end{bmatrix} \begin{bmatrix} \frac{1}{6} & \frac{3}{6} & \frac{5}{6} & \frac{1}{6} \\ \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{3}{6} \end{bmatrix} = \begin{bmatrix} \frac{26}{36} & -\frac{2}{36} & 0 & \frac{16}{36} \\ -\frac{2}{36} & \frac{10}{36} & \frac{16}{36} & 0 \\ 0 & \frac{16}{36} & \frac{26}{36} & \frac{2}{36} \\ \frac{16}{36} & 0 & \frac{2}{36} & \frac{10}{36} \end{bmatrix} \neq \mathbf{I}$$

This is not a contradiction because $\mathbf{C}$ is **not square**, and therefore a left-inverse (i.e. $\mathbf{C}^H$) need not (in fact will not) be a right-inverse. In fact, the inverse (i.e. $\mathbf{C}^{-1}$ is undefined for non-square matrices.)

4. Consider the matrix-vector equation

$$\mathbf{B}\,[\,\mathbf{x}\,]_{\mathbf{B}} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} [\,\mathbf{x}\,]_{\mathbf{B}} \;=\; \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix}$$

a. Solve this system for $[\,\mathbf{x}\,]_{\mathbf{B}}$ by Gaussian elimination.

solution:

For Gaussian elimination, first form the augmented matrix

$$\left[\begin{array}{ccc|c} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & 1 \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} & -7 \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} & 2 \end{array}\right]$$

and proceed

$$\begin{array}{c} \\ R_2 - 2R_1 \\ R_3 - 2R_1 \end{array} \quad \left[\begin{array}{ccc|c} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & 1 \\ 0 & -2 & 1 & -9 \\ 0 & -1 & 2 & 0 \end{array}\right]$$

Then

$$\begin{array}{c} \\ \\ R_3 - (1/2)R_2 \end{array} \quad \left[\begin{array}{ccc|c} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & 1 \\ 0 & -2 & 1 & -9 \\ 0 & 0 & \frac{3}{2} & \frac{9}{2} \end{array}\right]$$

Backsolving

$$c_3 = \frac{(9/2)}{(3/2)} = 3$$

$$c_2 = \frac{-9 - c_3}{(-2)} = \frac{-12}{-2} = 6$$

$$c_1 = \frac{1 - (2/3)c_2 + (2/3)c_3}{(1/3)} = \frac{1 - (2/3)(6) + (2/3)(3)}{(1/3)} = \frac{-1}{(1/3} = -3$$

solution:

and so

$$[\mathbf{x}]_{\mathbf{B}} = \begin{bmatrix} -3 \\ 6 \\ 3 \end{bmatrix}$$

b. Solve this system for $[\,\mathbf{x}\,]_{\mathbf{B}}$, without using elimination, but using the facts that $\mathbf{B}$ is unitary and $[\,\mathbf{x}\,]_{\mathbf{B}}$ represents the coordinates of $[\ 1\ \ -7\ \ \ 2\ ]^H$ in terms of the columns of $\mathbf{B}$.

solution:

Since we are given that $\mathbf{B}$ is unitary, then $\mathbf{B}^H\mathbf{B} = \mathbf{I}$, and so

$$\mathbf{B}\ [\mathbf{x}]_{\mathbf{B}} = \mathbf{x} \quad \Longrightarrow \quad [\mathbf{x}]_{\mathbf{B}} = \mathbf{B}^H\mathbf{x}$$

i.e., in this case, by direct computation

$$[\mathbf{x}]_{\mathbf{B}} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix} \begin{bmatrix} -3 \\ 6 \\ 3 \end{bmatrix}$$

which is (and must be for unitary matrices) precisely the same result as obtained by Gaussian elimination.

5. Show that if $\mathbf{Q}^{(1)}$ and $\mathbf{Q}^{(2)}$ are **any** two unitary matrices of the same size, then their product, i.e. $\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}$ is also Unitary.

solution:

By definition, the product $\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}$ will be unitary if and only if

$$\left(\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\right)^H \left(\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\right) = \mathbf{I}$$

But the Hermitian of a product is the product of the individual Hermitians, with the order reversed, i.e.

$$\left(\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\right)^H = \mathbf{Q}^{(2)^H}\mathbf{Q}^{(1)^H}$$

Therefore

$$\begin{aligned}\left(\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\right)^H \left(\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\right) &= \left(\mathbf{Q}^{(2)^H}\mathbf{Q}^{(1)^H}\right)\left(\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\right) \\ &= \mathbf{Q}^{(2)^H}\underbrace{\mathbf{Q}^{(1)^H}\mathbf{Q}^{(1)}}_{\mathbf{I}}\mathbf{Q}^{(2)} \\ &= \mathbf{Q}^{(2)^H}\mathbf{Q}^{(2)} = \mathbf{I}\end{aligned}$$

and so, by definition, $\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}$ is unitary.

6. Consider the three most common measures for the norm of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ (or the equivalent row or column vector forms):

$$\begin{aligned} \|\,\mathbf{x}\,\|_1 &= |x_1| + |x_2| + \cdots + |x_n| \\ \|\,\mathbf{x}\,\|_2 &= \sqrt{x_1^2 + x_2^2 + \cdots x_n^2} \\ \|\,\mathbf{x}\,\|_\infty &= \max_i |x_i| \end{aligned}$$

Compute each of these norms for each of the following vectors.

a. $\mathbf{x} = (-1, 2, -2)$

solution:

$$\begin{aligned} \|\,\mathbf{x}\,\|_1 &= |-1| + |2| + |-2| = 5 \\ \|\,\mathbf{x}\,\|_2 &= \sqrt{(-1)^2 + (2)^2 + (-2)^2} = \sqrt{9} = 3 \\ \|\,\mathbf{x}\,\|_\infty &= \max\{|-1|, |2|, |-2|\} = 2 \end{aligned}$$

b. $\mathbf{x} = (-4)$

solution:

$$\begin{aligned} \|\,\mathbf{x}\,\|_1 &= |-4| = 4 \\ \|\,\mathbf{x}\,\|_2 &= \sqrt{(-4)^2} = 4 \\ \|\,\mathbf{x}\,\|_\infty &= \max\{|-4|\} = 4 \end{aligned}$$

c. $\mathbf{x} = (10, -3, 12)$

solution:

$$\begin{aligned} \|\,\mathbf{x}\,\|_1 &= |10| + |-3| + |12| = 25 \\ \|\,\mathbf{x}\,\|_2 &= \sqrt{(10)^2 + (3)^2 + (12)^2} = \sqrt{253} = 15.9\ldots \\ \|\,\mathbf{x}\,\|_\infty &= \max\{|10|, |-3|, |12|\} = 12 \end{aligned}$$

d. $\mathbf{x} = (102, -17, -1)$

solution:

$$\begin{aligned} \|\,\mathbf{x}\,\|_1 &= |102| + |-17| + |-1| = 120 \\ \|\,\mathbf{x}\,\|_2 &= \sqrt{(102)^2 + (-17)^2 + (-1)^2} = \sqrt{10694} = 103.4\ldots \\ \|\,\mathbf{x}\,\|_\infty &= \max\{|102|, |-17|, |-1|\} = 102 \end{aligned}$$

e. $\mathbf{x} = (.1, -.2, -.4)$

solution:

$$\begin{aligned} \|\,\mathbf{x}\,\|_1 &= |.1| + |-.2| + |-.4| = .7 \\ \|\,\mathbf{x}\,\|_2 &= \sqrt{(.1)^2 + (-.2)^2 + (-.4)^2} = \sqrt{0.21} = 0.458\ldots \\ \|\,\mathbf{x}\,\|_\infty &= \max\{|.1|, |-.2|, |-.4|\} = .4 \end{aligned}$$

f. $\mathbf{x} = (-12, -2, 4, 6, 5)$

solution:

$$\begin{aligned} \|\,\mathbf{x}\,\|_1 &= |-12| + |-2| + |4||6||5| = 29 \\ \|\,\mathbf{x}\,\|_2 &= \sqrt{(12)^2 + (-2)^2 + (4)^2 + (6)^2 + (5)^2} = \sqrt{225} = 15 \\ \|\,\mathbf{x}\,\|_\infty &= \max\{|-12|, |-2|, |-4|, |6|, |5|\} = 12 \end{aligned}$$

Note that, in all of these cases,

$$\|\,\mathbf{x}\,\|_\infty \leq \|\,\mathbf{x}\,\|_2 \leq \|\,\mathbf{x}\,\|_1$$

7. Any vector norm *induces* a corresponding matrix norm according to the relationship:

$$\| \mathbf{A} \| = \max_{\mathbf{x}} \frac{\| \mathbf{A}\,\mathbf{x} \|}{\| \mathbf{x} \|}$$

Show that, for any matrix norm and any scalar α, $\quad \| \alpha\, \mathbf{A} \| = |\alpha|\, \| \mathbf{A} \|$

solution:

We know that for each of the *vector* norms, from which the matrix norms are derived, that

$$\| \alpha\mathbf{x} \| = |\alpha|\, \| \mathbf{x} \| \quad , \quad \text{e.g.} \quad \| \alpha\mathbf{x} \|_1 = |\alpha|\, \| \mathbf{x} \|_1$$

since that is simply one of the basic "rules" which any norm must satisfy in order to be a norm. But, because $(\alpha\mathbf{A})\,\mathbf{x} = \alpha\,(\mathbf{A}\,\mathbf{x})$ for it all matrices, and because $\mathbf{A}\,\mathbf{x}$ is also a vector, then, in any norm,

$$\| (\alpha\mathbf{A})\,\mathbf{x} \| = \| \alpha\,(\mathbf{A}\,\mathbf{x}) \| = |\alpha| \| \mathbf{A}\,\mathbf{x} \|$$

Therefore, by definition:

$$\| \alpha\mathbf{A} \| = \max_{\mathbf{x}} \frac{\| (\alpha\mathbf{A})\ \mathbf{x} \|}{\| \mathbf{x} \|} = \max_{\mathbf{x}} \frac{|\alpha| \| \mathbf{A}\,\mathbf{x} \|}{\| \mathbf{x} \|} = |\alpha| \max_{\mathbf{x}} \frac{\| \mathbf{A}\,\mathbf{x} \|}{\| \mathbf{x} \|} \ \equiv\ |\alpha| \| \mathbf{A} \|$$

8. Consider the three matrices:

$$\text{a.}\quad \begin{bmatrix} 2 & -5 & 1 \\ 1 & 3 & 2 \end{bmatrix} \qquad \text{b.}\quad \begin{bmatrix} 10 & 7 & -2 \\ 6 & 4 & -1 \\ -2 & 1 & 1 \end{bmatrix} \qquad \text{c.}\quad \begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 1 & 2 \end{bmatrix}$$

For each of these matrices, and for each of the norms described in problem 6:

(i.) Pick five different "input" vectors ($\mathbf{x}$). (Make sure at least some of them have some *negative* components!)

(ii.) For each of these inputs, compute the ratio $\dfrac{\| \mathbf{A}\,\mathbf{x} \|}{\| \mathbf{x} \|}$

(iii.) Based on your answers to part b., determine a lower bound for $\| \mathbf{A} \|$

(iv.) Compare your lower bound to the actual corresponding value of $\| \mathbf{A} \|$ as determined by MATLAB's relevant **norm( )** command.

solution:

a. For $\begin{bmatrix} 2 & -5 & 1 \\ 1 & 3 & 2 \end{bmatrix}$, pick as inputs

$$\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} ,\ \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} ,\ \mathbf{x}^{(3)} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} ,\ \mathbf{x}^{(4)} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} ,\ \mathbf{x}^{(5)} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

Then $\quad \mathbf{A}\,\mathbf{x}^{(1)} = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$ and so

$$\frac{\| \mathbf{A}\,\mathbf{x}^{(1)} \|_1}{\| \mathbf{x}^{(1)} \|_1} = \frac{8}{3} \doteq 2.67 \,, \quad \frac{\| \mathbf{A}\,\mathbf{x}^{(1)} \|_2}{\| \mathbf{x}^{(1)} \|_2} = \frac{\sqrt{40}}{\sqrt{3}} \doteq 3.65 \,, \quad \frac{\| \mathbf{A}\,\mathbf{x}^{(1)} \|_\infty}{\| \mathbf{x}^{(1)} \|_\infty} = \frac{6}{1} = 6$$

Continuing, we compute $\ \mathbf{A}\,\mathbf{x}^{(2)} = \begin{bmatrix} 8 \\ 0 \end{bmatrix}$ and so

$$\frac{\| \mathbf{A}\,\mathbf{x}^{(2)} \|_1}{\| \mathbf{x}^{(2)} \|_1} = \frac{8}{3} \doteq 2.67 \,, \quad \frac{\| \mathbf{A}\,\mathbf{x}^{(2)} \|_2}{\| \mathbf{x}^{(2)} \|_2} = \frac{8}{\sqrt{3}} \doteq 4.62 \,, \quad \frac{\| \mathbf{A}\,\mathbf{x}^{(2)} \|_\infty}{\| \mathbf{x}^{(2)} \|_\infty} = \frac{8}{1} = 8$$

Similarly, $\quad \mathbf{A}\,\mathbf{x}^{(3)} = \begin{bmatrix} -5 \\ 10 \end{bmatrix}$, $\quad$ and so

$$\frac{\| \mathbf{A}\,\mathbf{x}^{(3)} \|_1}{\| \mathbf{x}^{(3)} \|_1} = \frac{15}{5} = 3 \,, \quad \frac{\| \mathbf{A}\,\mathbf{x}^{(3)} \|_2}{\| \mathbf{x}^{(3)} \|_2} = \frac{\sqrt{125}}{\sqrt{9}} \doteq 3.73 \,, \quad \frac{\| \mathbf{A}\,\mathbf{x}^{(3)} \|_\infty}{\| \mathbf{x}^{(3)} \|_\infty} = \frac{10}{2} = 5$$

solution:

$\mathbf{A}\,\mathbf{x}^{(4)} = \begin{bmatrix} 15 \\ -2 \end{bmatrix}$ and so

$$\frac{\|\,\mathbf{A}\,\mathbf{x}^{(4)}\,\|_1}{\|\,\mathbf{x}^{(4)}\,\|_1} = \frac{17}{5} = 3.4\,, \quad \frac{\|\,\mathbf{A}\,\mathbf{x}^{(4)}\,\|_2}{\|\,\mathbf{x}^{(4)}\,\|_2} = \frac{\sqrt{229}}{\sqrt{9}} \doteq 5.04\,, \quad \frac{\|\,\mathbf{A}\,\mathbf{x}^{(4)}\,\|_\infty}{\|\,\mathbf{x}^{(4)}\,\|_\infty} = \frac{15}{2} = 7.5$$

and finally $\quad \mathbf{A}\,\mathbf{x}^{(5)} = \begin{bmatrix} 9 \\ 2 \end{bmatrix} \quad$ and so

$$\frac{\|\,\mathbf{A}\,\mathbf{x}^{(5)}\,\|_1}{\|\,\mathbf{x}^{(5)}\,\|_1} = \frac{11}{4} = 2.75\,, \quad \frac{\|\,\mathbf{A}\,\mathbf{x}^{(5)}\,\|_2}{\|\,\mathbf{x}^{(5)}\,\|_2} = \frac{\sqrt{85}}{\sqrt{6}} \doteq 3.76\,, \quad \frac{\|\,\mathbf{A}\,\mathbf{x}^{(5)}\,\|_\infty}{\|\,\mathbf{x}^{(5)}\,\|_\infty} = \frac{9}{2} = 4.5$$

The results of the above can be summarized in the following table:

	$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	$\mathbf{x}^{(3)}$	$\mathbf{x}^{(4)}$	$\mathbf{x}^{(5)}$	Max.	MATLAB
$\frac{\|\,\mathbf{A}\,\mathbf{x}^{(i)}\,\|_1}{\|\,\mathbf{x}^{(i)}\,\|_1}$	2.67	2.67	3	3.4	2.75	3.4	8
$\frac{\|\,\mathbf{A}\,\mathbf{x}^{(i)}\,\|_2}{\|\,\mathbf{x}^{(i)}\,\|_2}$	3.65	4.62	3.73	5.04	3.76	5.04	5.97
$\frac{\|\,\mathbf{A}\,\mathbf{x}^{(i)}\,\|_\infty}{\|\,\mathbf{x}^{(i)}\,\|_\infty}$	6	8	5	7.5	4.5	8	8

---

b. For $\begin{bmatrix} 10 & 7 & -2 \\ 6 & 4 & -1 \\ -2 & 1 & 1 \end{bmatrix}$, pick as inputs

$$\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \ \mathbf{x}^{(3)} = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \ \mathbf{x}^{(4)} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \ \mathbf{x}^{(5)} = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$$

Then $\quad \mathbf{A}\,\mathbf{x}^{(1)} = \begin{bmatrix} 15 \\ 9 \\ 0 \end{bmatrix}, \ \mathbf{A}\,\mathbf{x}^{(2)} = \begin{bmatrix} 19 \\ 11 \\ -2 \end{bmatrix}, \ \mathbf{A}\,\mathbf{x}^{(3)} = \begin{bmatrix} 36 \\ 21 \\ -3 \end{bmatrix},$

$\mathbf{A}\,\mathbf{x}^{(4)} = \begin{bmatrix} -7 \\ -4 \\ 5 \end{bmatrix}$, and $\mathbf{A}\,\mathbf{x}^{(5)} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

solution:

Performing similar computations to those in part a. yields the following table:

	$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	$\mathbf{x}^{(3)}$	$\mathbf{x}^{(4)}$	$\mathbf{x}^{(5)}$	Max.	MATLAB
$\frac{\| \mathbf{A}\,\mathbf{x}^{(i)} \|_1}{\| \mathbf{x}^{(i)} \|_1}$	8	10.67	12	4	.50	12	18
$\frac{\| \mathbf{A}\,\mathbf{x}^{(i)} \|_2}{\| \mathbf{x}^{(i)} \|_2}$	10.1	12.7	13.9	3.87	.58	13.9	14.4
$\frac{\| \mathbf{A}\,\mathbf{x}^{(i)} \|_\infty}{\| \mathbf{x}^{(i)} \|_\infty}$	15	19	18	3.5	.5	19	19

---

b. For $\begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 1 & 2 \end{bmatrix}$, pick as inputs

$$\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} , \ \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} , \ \mathbf{x}^{(3)} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} , \ \mathbf{x}^{(4)} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} , \ \mathbf{x}^{(5)} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Then $\mathbf{A}\,\mathbf{x}^{(1)} = \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{A}\,\mathbf{x}^{(2)} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{A}\,\mathbf{x}^{(3)} = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}$,

$\mathbf{A}\,\mathbf{x}^{(4)} = \begin{bmatrix} -3 \\ 2 \\ 3 \end{bmatrix}$, and $\mathbf{A}\,\mathbf{x}^{(5)} = \begin{bmatrix} -1 \\ 6 \\ 5 \end{bmatrix}$

Again performing similar computations to those in part a. yields the following table:

	$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	$\mathbf{x}^{(3)}$	$\mathbf{x}^{(4)}$	$\mathbf{x}^{(5)}$	Max.	MATLAB
$\frac{\| \mathbf{A}\,\mathbf{x}^{(i)} \|_1}{\| \mathbf{x}^{(i)} \|_1}$	3.5	1.5	1.67	2.67	4	4	5
$\frac{\| \mathbf{A}\,\mathbf{x}^{(i)} \|_2}{\| \mathbf{x}^{(i)} \|_2}$	3.54	1.58	1.61	2.10	3.52	3.52	3.57
$\frac{\| \mathbf{A}\,\mathbf{x}^{(i)} \|_\infty}{\| \mathbf{x}^{(i)} \|_\infty}$	4	2	1.50	1.50	3	4	4

9. The infinity norm of a vector:

$$\| x \|_\infty = \max_i |x_i|$$

i.e., the component with the greatest magnitude, induces a corresponding matrix norm:

$$\| \mathbf{A} \|_\infty = \max_{\mathbf{x}} \frac{\| \mathbf{A}\,\mathbf{x} \|_\infty}{\| \mathbf{x} \|_\infty} = \max_{\| \mathbf{x} \|_\infty = 1} \| \mathbf{A}\,\mathbf{x} \|_\infty$$

Show that $\| \mathbf{A} \|_\infty = \max_i \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}$, or, in other words, the matrix infinity norm is just the largest of the sums of the magnitudes of the coefficients on each row. For this reason, the matrix infinity norm is commonly called the *row-sum norm.*

solution:

First note that because the vector infinity norm is based on the largest component,

$$\| \mathbf{x} \|_\infty = \max_j |x_j| = 1 \quad \Longrightarrow \quad |x_j| \leq 1 \quad , \quad \text{all} \ \ j$$

Therefore, because the absolute value of a sum can be no greater than the sum of the absolute values, and because no single component of any vector $\mathbf{x}$ of interest here can exceed one in magnitude

$$|(\mathbf{A}\,\mathbf{x})_i| = \left| \sum_{j=1}^{n} A_{ij} x_j \right| \leq \sum_{j=1}^{n} |A_{ij} x_j| = \sum_{j=1}^{n} |A_{ij}| \cdot |x_j| \leq \sum_{j=1}^{n} |A_{ij}| \quad , \quad \text{all} \ \ i$$

Therefore $$\max_i \{ |(\mathbf{A}\,\mathbf{x})_i| \} \leq \max_i \left\{ \sum_{j=1}^{n} |A_{ij}| \right\}$$

Finally, noting that the right hand side of this last equation is independent of the components of $\mathbf{x}$ (provided they are all less than or equal to one in magnitude), we conclude

$$\| \mathbf{A} \|_\infty \equiv \max_{\| \mathbf{x} \|_\infty = 1} \left[\max_i \{ |(\mathbf{A}\,\mathbf{x})_i| \} \right] \leq \max_i \left\{ \sum_{j=1}^{n} |A_{ij}| \right\}$$

solution:

Now, for any particular matrix $\mathbf{A}$, let i_{max} denote the row on which this maximum row sum occurs. Then define the (input) vector $\tilde{\mathbf{x}}$ by

$$\tilde{x}_j = \text{sign } [A_{i_{max}j}] = \begin{cases} +1 & , \ A_{i_{max}j} \geq 0 \\ -1 & , \ A_{i_{max}j} < 0 \end{cases}$$

Observe that $\| \tilde{\mathbf{x}} \|_\infty = 1$, and for every j

$$A_{i_{max}j}\, \tilde{x}_j = \text{sign } [A_{i_{max}j}] \cdot A_{i_{max}j} = |A_{i_{max}j}|$$

Therefore, by the way we constructed i_{max} and $\tilde{\mathbf{x}}$,

$$\sum_{j=1}^{n} A_{i_{max}j}\, \tilde{x}_j = \sum_{j=1}^{n} |A_{i_{max}j}| = \max_i \left\{ \sum_{j=1}^{n} |A_{ij}| \right\} = \max_i \left\{ \left| \sum_{j=1}^{n} A_{ij} \tilde{x}_j \right| \right\}$$

(The last equality occurs because $|\tilde{x}_j| = 1$ for all j, and therefore, if a larger sum were possible on any other row, it would imply we had chosen i_{max} incorrectly.)

But now, because $\tilde{\mathbf{x}}$ is only one of the vectors for which $\| \mathbf{x} \|_\infty = 1$,

$$\| \mathbf{A} \|_\infty \equiv \max_{\| \mathbf{x} \|_\infty = 1} \left[\max_i \{ |(\mathbf{A}\,\mathbf{x})_i| \} \right] = \max_{\| \mathbf{x} \|_\infty = 1} \left[\max_i \left\{ \left| \sum_{j=1}^{n} A_{ij}\, x_j \right| \right\} \right]$$

$$\geq \max_i \left\{ \left| \sum_{j=1}^{n} A_{ij}\, \tilde{x}_j \right| \right\} = \max_i \left\{ \sum_{j=1}^{n} |A_{ij}| \right\}$$

However, at this point, we have shown that **both**

$$\| \mathbf{A} \|_\infty \leq \max_i \left\{ \sum_{j=1}^{n} |A_{ij}| \right\} \quad \textbf{and} \quad \| \mathbf{A} \|_\infty \geq \max_i \left\{ \sum_{j=1}^{n} |A_{ij}| \right\}$$

The only possible conclusion is: $\| \mathbf{A} \|_\infty = \max_i \left\{ \sum_{j=1}^{n} |A_{ij}| \right\}$.

10. Consider the following vectors and matrices:

$$\text{(a.)}\ \mathbf{x} = \begin{bmatrix} -6 \\ 4 \\ -4 \end{bmatrix} \qquad \text{(b.)}\ \mathbf{y} = \begin{bmatrix} -9 \\ 2 \\ -9 \\ -2 \end{bmatrix} \qquad \text{(c.)}\ \mathbf{z} = \begin{bmatrix} -1 \\ -4 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

$$\text{(d.)}\ \mathbf{A} = \begin{bmatrix} 8 & 8 & -4 & 2 \\ 5 & -6 & -1 & -2 \\ -7 & -4 & -9 & 0 \\ -10 & 3 & 10 & -3 \end{bmatrix} \qquad \text{(e.)}\ \mathbf{B} = \begin{bmatrix} -6 & -10 & -2 & 7 & 0 \\ -6 & 5 & 7 & -10 & 4 \\ 2 & -1 & 1 & 4 & -1 \\ -5 & 9 & -6 & -2 & -4 \\ -6 & -1 & 3 & 7 & -6 \end{bmatrix}$$

Find the infinity norm of each.

solution:

For parts (a.)-(c.), the infinity norm is simply the coordinate with the largest absolute value. Therefore

$$\|\,\mathbf{x}\,\|_\infty = |-6| = 6 \quad , \quad \|\,\mathbf{y}\,\|_\infty = |-9| = 9 \quad , \quad \text{and} \quad \|\,\mathbf{z}\,\|_\infty = |5| = 5$$

For parts (d.) and (e.), the infinity norm is the maximum of the sum of the absolute values of the elements on each row. Therefore:

$$\mathbf{A} = \begin{bmatrix} 8 & 8 & -4 & 2 \\ 5 & -6 & -1 & -2 \\ -7 & -4 & -9 & 0 \\ -10 & 3 & 10 & -3 \end{bmatrix} \quad \begin{array}{l} \rightarrow \ \sum |a_{1j}| = 22 \\ \rightarrow \ \sum |a_{2j}| = 14 \\ \rightarrow \ \sum |a_{3j}| = 20 \\ \rightarrow \ \sum |a_{4j}| = 26 \end{array}$$

and so $\|\,A\,\|_\infty = 26$. Similarly

$$\mathbf{B} = \begin{bmatrix} -6 & -10 & -2 & 7 & 0 \\ -6 & 5 & 7 & -10 & 4 \\ 2 & -1 & 1 & 4 & -1 \\ -5 & 9 & -6 & -2 & -4 \\ -6 & -1 & 3 & 7 & -6 \end{bmatrix} \quad \begin{array}{l} \rightarrow \ \sum |a_{1j}| = 25 \\ \rightarrow \ \sum |a_{2j}| = 32 \\ \rightarrow \ \sum |a_{2j}| = \ \ 9 \\ \rightarrow \ \sum |a_{3j}| = 26 \\ \rightarrow \ \sum |a_{4j}| = 23 \end{array}$$

and so $\|\,B\,\|_\infty = 32$.

11. Show directly that the one norm, defined by

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|$$

satisfies the following general properties:

(a.) $\|\mathbf{x}\|_1 > 0\ ,\quad \mathbf{x} \neq 0$

solution:

By definition of the absolute value, $|x_i| \geq 0$ for all x_i. Therefore, the sum of any number of absolute values cannot be negative. Now suppose $\mathbf{x} \neq \mathbf{0}$. Then $\mathbf{x}$ must have at least one non-zero component. We can assume, without loss of generality, that it's x_1. So, $|x_1| > 0$. Then

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| = \underbrace{|x_1|}_{>0} + \underbrace{|x_2| + \cdots + |x_n|}_{\geq 0} \geq |x_1| > 0$$

(b.) $\|\mathbf{x}+\mathbf{y}\|_1 \leq \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$, where $\mathbf{x}$ and $\mathbf{y}$ are any vectors. (This so-called *triangle inequality* essentially ensures the shortest "distance" between points must be the line connecting them.)

solution:

A basic property of the absolute value is $|\alpha + \beta| \leq |\alpha| + |\beta|$, for any real values α and β. Therefore, since each component of $\mathbf{x}+\mathbf{y}$ is precisely the sum of the corresponding components of $\mathbf{x}$ and $\mathbf{y}$:

$$\|\mathbf{x}+\mathbf{y}\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leq \sum_{i=1}^{n} [\,|x_i| + |y_i|\,] = \underbrace{\sum_{i=1}^{n} |x_i|}_{\equiv \|\mathbf{x}\|_1} + \underbrace{\sum_{i=1}^{n} |y_i|}_{\equiv \|\mathbf{y}\|_1}$$

or, in summary:

$$\|\mathbf{x}+\mathbf{y}\|_1 \leq \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$$

(c.) $\|\alpha\mathbf{x}\|_1 = |\alpha|\|\mathbf{x}\|_1$, where $\mathbf{x}$ is any vector and α is any scalar. (This formula ensures that multiplying any vector by a scalar factor simply changes its "length" by that factor, and also that $\|\mathbf{0}\|_1 = 0$.)

solution:

By definition, the i^{th} component of $\alpha\mathbf{x}$ is precisely αx_i, where x_i is the corresponding component of $\mathbf{x}$. Therefore:

$$\| \alpha\mathbf{x} \|_1 = \sum_{i=1}^{n} |\alpha x_i| = \sum_{i=1}^{n} |\alpha||x_i| = |\alpha| \left\{ \sum_{i=1}^{n} |x_i| \right\} = \alpha \| \mathbf{x} \|_1$$

12. Using MATLAB, graph

$$\left\{ \mathbf{A}\,\mathbf{x} \mid x_1^2 + x_2^2 = 1 \right\}$$

for each of the following matrices. Based on your figures, estimate the singular values of each matrix, and then compare your estimates to the results of MATLAB's **svd( )** command.

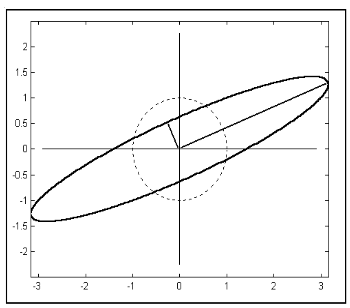
a. $\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$

solution:

One example of MATLAB code that can generate these plots is simply

```
a= [ 3, 1 ; 1,  1 ] ;
theta = 0:0.01:2*pi ;
x = cos(theta) ; y = sin(theta) ;
helipse = a*[ x; y ] ;
plot( x, y, helipse(1,:), helipse(2,:), '- -') ;
axis equal ;
```

where the last command is necessary to ensure the cirle of radius one looks like a circle. In this case, the resulting figure looks like:



where the semimajor and semiminor axes of the resulting ellipse have been added. From this figure, it appears that these axes are of lengths

$$\sigma_1 \doteq 3.5 \quad \text{and} \quad \sigma_2 \doteq .6$$

These "eyeball" values compare favorably with the results returned by MATLAB, i.e.

$$\sigma_1 = 3.4142 \quad \text{and} \quad \sigma_2 = 0.5858$$

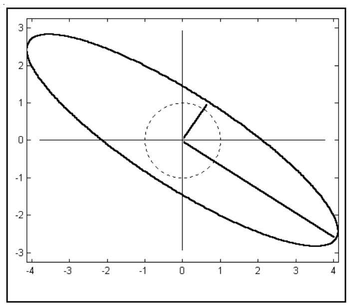
b. $\begin{bmatrix} -4 & -1 \\ 2 & 2 \end{bmatrix}$

solution:

We can use the identical MATLAB code from part a., except, of course, changing the entries in the matrix, i.e. :

a= [-4, -1 ; 2, 2] ;

(Note the commas between elements are optional, but recommended in the case of negative entries.) In this case then, the resulting figure looks like:



where the semimajor and semiminor axes of the resulting ellipse have again been added. From this figure, it appears that, in this case, these axes are of lengths

$$\sigma_1 \doteq 5.0 \quad \text{and} \quad \sigma_2 \doteq 1.3$$

These "eyeball" values again compare favorably with the results returned by MATLAB, i.e.

$$\sigma_1 = 4.8442 \quad \text{and} \quad \sigma_2 = 1.2386$$

c. $\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$

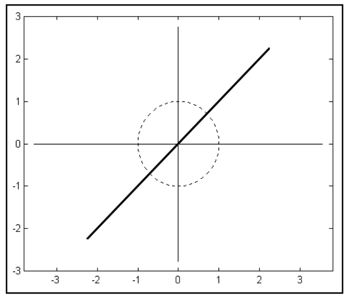
solution:

We can again use the identical MATLAB code from part a., except, of course, again changing the entries in the matrix, i.e. :

a= [2 1 ; 2 1] ;

solution:

In this case then, the resulting figure looks like:



In this case, however the ellipse has degenerated to simply a straight line segment. This is a clear indication of at least one zero (or near zero) singular value. (The other singular value here will be the half-length of the line segment.) Therefore, from this figure, it appears that the singular values here are

$$\sigma_1 \doteq 3.1 \quad \text{and} \quad \sigma_2 \doteq 0$$

These "eyeball" values again compare favorably with the results returned by MATLAB, i.e.

$$\sigma_1 = 3.1623 \quad \text{and} \quad \sigma_2 = 0$$

13. Find the singular value decomposition of each of the following matrices, and compare your results to the results of MATLAB's **svd( )** command.

a. $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

solution:

In general, finding the SVD of a matrix is an extremely difficult computational task. However, there are certain exceptions, and this is one. Specifically, in this case, because $\mathbf{A}$ is already diagonal:

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{I}$$

which is, by definition, an SVD of $\mathbf{A}$, with $\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^H = \mathbf{I}$.

The MATLAB command **[u s v] = svd(a)** produces exactly the same SVD.

b. $\begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}$

solution:

This case is virtually identical to part a., except for the negative entry on the diagonal, because, by definition, singular values must be non-negative. That problem, however is very easily remidied, since we can just write:

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} = \mathbf{I} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

which is, again, by definition, an SVD of $\mathbf{A}$, with $\mathbf{U} = \mathbf{I}$ and

$$\mathbf{V}^H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{V} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{V}^H \mathbf{V} = \mathbf{I}$$

and so $\mathbf{V}$ is unitary.

solution:

The MATLAB command **[u s v] = svd(a)** produces a slight variant of this SVD, i.e.

$$\mathbf{u} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad , \quad \mathbf{s} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(Such a situation, i.e. agreement except for plus and minus changes, are quite common in numerical linear algebra.)

c. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

solution:

In this case, the matrix is not diagonal, so we can't use the "trick" from part a. or b. However, note that the columns of the matrix are identical, and therefore, trivially,

$$\text{Col}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad \text{Null}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

But we do know that, in general, Col($\mathbf{A}$) is the span of the columns of $\mathbf{U}$ corresponding to the non-zero singular values of $\mathbf{A}$, and Null($\mathbf{A}$) is the span of the columns of $\mathbf{V}$ corresponding to the zero singular values. So the columns of $\mathbf{U}$ and $\mathbf{V}$ corresponding, respectively, to the non-zero and zero singular values of $\mathbf{A}$ are (except for possible sign reversals):

$$\mathbf{u}^{(1)} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad \mathbf{v}^{(2)} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Since we are working in $\mathbb{R}^2$, and we know that $\mathbf{U}$ and $\mathbf{V}$ are unitary and therefore their columns are orthonormal bases for $\mathbb{R}^2$, then we can easily extend the existing columns above to the full matrices

$$\mathbf{U} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad \mathbf{V} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

solution:

(Obviously, $\mathbf{U} = \mathbf{V}$ in this case. That actually, should not be surprising, based on the fact $\mathbf{A}$ is symmetric.) Direct computation will now show that

$$\mathbf{U}^H\mathbf{A}\mathbf{V} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \equiv \boldsymbol{\Sigma}$$

which is, of course, equivalent to $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$, i.e. an SVD of $\mathbf{A}$.

The MATLAB command **[u s v] = svd(a)** produces a slight variant of this SVD, i.e.

$$\mathbf{u} = \begin{bmatrix} 0.7071 & -0.7071 \\ 0.7071 & 0.7071 \end{bmatrix} \quad , \quad \mathbf{s} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 0.7071 & -0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}$$

which, again, is identical to our result except for the not uncommon sign change.

14. Using the singular value decomposition, show that, for any matrix $\mathbf{A}$,

$$\text{Null}(\mathbf{A^H A}) = \text{Null}(\mathbf{A})$$

solution:

We know that any matrix $\mathbf{A}$ has a **full** singular value decomposition, i.e.

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H \quad \text{where} \quad \mathbf{U}^H\mathbf{U} = \mathbf{I} \ , \quad \mathbf{V}^H\mathbf{V} = \mathbf{I} \ ,$$

$\mathbf{U}$ and $\mathbf{V}$ are square, and $\boldsymbol{\Sigma}$ is diagonal, with only real, non-negative entries. But then

$$\mathbf{A}^H\mathbf{A} = \left(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H\right)^H \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H = \left(\mathbf{V}\boldsymbol{\Sigma}^H\mathbf{U}^H\right)\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H = \mathbf{V}\boldsymbol{\Sigma}^H \underbrace{\left(\mathbf{U}^H\mathbf{U}\right)}_{\mathbf{I}} \boldsymbol{\Sigma}\mathbf{V}^H$$

or

$$\mathbf{A}^H\mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}^2\mathbf{V}^H$$

where $\boldsymbol{\Sigma}^2$ is the diagonal matrix whose entries are the squares of the singular values of $\mathbf{A}$. But, by definition, because $\mathbf{V}$ is unitary, $\mathbf{V}\boldsymbol{\Sigma}^2\mathbf{V}^H$ is a singular value decomposition of $\mathbf{A}^H\mathbf{A}$. Moreover, the null space of $\mathbf{A}^H\mathbf{A}$ is known to be precisely the span of the columns of $\mathbf{V}$ corresponding to the zero diagonal elements of $\boldsymbol{\Sigma}^2$. Therefore, since the diagonal elements of $\boldsymbol{\Sigma}^2$ are exactly the squares of the diagonal elements of $\boldsymbol{\Sigma}$, then the zero diagonal elements of $\boldsymbol{\Sigma}$ are identical with the zero diagonal elements of $\boldsymbol{\Sigma}$, and so the columns of $\mathbf{V}$ corresponding to the zero diagonal elements of $\boldsymbol{\Sigma}^2$ are identical with the columns of $\mathbf{V}$ corresponding to the zero diagonal elements of $\boldsymbol{\Sigma}$. Hence their spans are identical, and so

$$\text{Null}(\mathbf{A}^H\mathbf{A}) = \text{Null}(\mathbf{A})$$

15. The SVD is commonly developed from an eigenvalue and eigenvector approach instead of from the viewpoint of Euclidean norms. Specifically, we can show that, for any matrix $\mathbf{A}$, the singular values satisfy $\sigma_i = \sqrt{|\lambda_i|}$, where the λ_i are the eigenvalues of $\mathbf{A}^H\mathbf{A}$, and the right singular vectors of $\mathbf{A}$ (i.e., the $\mathbf{v}^{(i)}$) are the eigenvectors of $\mathbf{A}^H\mathbf{A}$. This leads to constructing the (reduced) SVD via the following procedure:

(i.) Form the product $\mathbf{A}^H\,\mathbf{A}$.

(ii.) Find the eigenvalues (λ_i) and eigenvectors $(\mathbf{v}^{(i)})$ of that product.

(iii.) For each $\lambda_i \neq 0$, define $\sigma_i = \sqrt{\lambda_i}$, where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > 0$.

(iv.) For each $\sigma_i \neq 0$, define $\mathbf{u}^{(i)} = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}^{(i)}$, where $\mathbf{v}^{(i)}$ is the eigenvector associated with λ_i.

(Unfortunately, as we shall see later, this procedure is not particularly well suited for "large" matrices.) Apply this procedure to find the SVD of the matrix

$$\mathbf{A} = \begin{bmatrix} 26 & 18 \\ 1 & 18 \\ 14 & 27 \end{bmatrix}$$

solution:

Proceeding in order:

(i.) Form the product $\mathbf{A}^H\,\mathbf{A}$.

$$\mathbf{A}^H\,\mathbf{A} = \begin{bmatrix} 26 & 1 & 14 \\ 18 & 18 & 27 \end{bmatrix} \begin{bmatrix} 26 & 18 \\ 1 & 18 \\ 14 & 27 \end{bmatrix} = \begin{bmatrix} 873 & 864 \\ 864 & 1377 \end{bmatrix}$$

(ii.) Find the eigenvalues (λ_i) and eigenvectors $(\mathbf{v}^{(i)})$ of that product.

The characteristic equation of the product is

$$\det\left(\mathbf{A}^H\,\mathbf{A} - \lambda\mathbf{I}\right) = \det\begin{bmatrix} (873-\lambda) & 864 \\ 864 & (1377-\lambda) \end{bmatrix}$$

$$= \lambda^2 - 2250\lambda + 455625 = 0$$

$$\Longrightarrow \quad \lambda = 2025\ ,\ 225$$

Then, for $\lambda_1 = 2025$, $\left(\mathbf{A}^H\,\mathbf{A} - \lambda_1\mathbf{I}\right)\mathbf{v}^{(1)} = \mathbf{0}$ leads to the augmented matrix:

$$\left[\begin{array}{cc|c} -1152 & 864 & 0 \\ 864 & -648 & 0 \end{array}\right] \quad \Longrightarrow \quad \mathbf{v}^{(1)} = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$$

where we have normalized the eigenvector $\mathbf{v}^{(1)}$.

solution:

Similarly, for Then, for $\lambda_2 = 225$, $\left(\mathbf{A}^H\,\mathbf{A} - \lambda_2\mathbf{I}\right)\mathbf{v}^{(2)} = \mathbf{0}$ leads to the augmented matrix:

$$\left[\begin{array}{cc:c} 648 & 864 & 0 \\ 864 & 1152 & 0 \end{array}\right] \quad \Longrightarrow \quad \mathbf{v}^{(2)} = \begin{bmatrix} \frac{4}{5} \\ -\frac{3}{5} \end{bmatrix}$$

where we have again normalized the eigenvector $\mathbf{v}^{(2)}$. Therefore, for the SVD,

$$\mathbf{V} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}$$

where the columns of $\mathbf{V}$ are obviously orthonormal. (Note also that $\mathbf{V}$ is symmetric, and so, in this case, we will also have $\mathbf{V}^H = \mathbf{V}$. Note also $\mathbf{V}$ is square, and therefore unitary. So we omit the "hat" that we use to signify reduced matrices.)

(iii.) For each $\lambda_i \neq 0$, define $\sigma_i = \sqrt{\lambda_i}$, where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > 0$.

So

$$\lambda_1 = 2025 \quad \Longrightarrow \quad \sigma_1 = 45 \quad \text{and} \quad \lambda_2 = 225 \quad \Longrightarrow \quad \sigma_2 = 15$$

and so

$$\boldsymbol{\Sigma} = \begin{bmatrix} 45 & 0 \\ 0 & 15 \end{bmatrix}$$

(iv.) For each $\sigma_i \neq 0$, define $\mathbf{u}^{(i)} = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}^{(i)}$, where $\mathbf{v}^{(i)}$ is the eigenvector associated with λ_i.

So, for $\sigma_1 = 45$, we have

$$\mathbf{u}^{(1)} = \frac{1}{\sigma_1}\,\mathbf{A}\,\mathbf{v}^{(1)} \quad \Longrightarrow \quad \mathbf{u}^{(i)} = \frac{1}{45}\begin{bmatrix} 26 & 18 \\ 1 & 18 \\ 14 & 27 \end{bmatrix}\begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

(Note, we must use $\mathbf{A}$ here, **not** $\mathbf{A}^H\,\mathbf{A}$.)

solution:

Similarly, or $\sigma_2 = 15$,

$$\mathbf{u}^{(2)} = \frac{1}{\sigma_2}\,\mathbf{A}\,\mathbf{v}^{(2)} \quad \Longrightarrow \quad \mathbf{u}^{(i)} = \frac{1}{15}\begin{bmatrix} 26 & 18 \\ 1 & 18 \\ 14 & 27 \end{bmatrix}\begin{bmatrix} \frac{4}{5} \\ -\frac{3}{5} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ -\frac{1}{3} \end{bmatrix}$$

Therefore, combining these two, we have

$$\hat{\mathbf{U}} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

(Note we now use the "hat" since this $\mathbf{U}$ is not square, and so we are generating a true reduced SVD. This also means we should have used $\hat{\boldsymbol{\Sigma}}$ earlier as well.)

Direct computation will verify that

$$\hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}\,\mathbf{V}^H = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}\begin{bmatrix} 45 & 0 \\ 0 & 15 \end{bmatrix}\begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix} = \begin{bmatrix} 26 & 18 \\ 1 & 18 \\ 14 & 27 \end{bmatrix} = \mathbf{A}$$

Furthermore, we could, if we chose, show that the full SVD of $\mathbf{A}$ was:

$$\mathbf{A} = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \end{bmatrix}\begin{bmatrix} 45 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}$$

16. Repeat the calculation of the (reduced) SVD using the eigenvalue and eigenvector approach from problem 15 for the matrix:

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 2 & 4 \\ 2 & 4 \\ 2 & 4 \end{bmatrix}$$

solution:

Proceeding in order:
(i.) Form the product $\mathbf{A}^H\,\mathbf{A}$.

$$\mathbf{A}^H\,\mathbf{A} = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 4 & 4 & 4 & 4 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 2 & 4 \\ 2 & 4 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 16 & 32 \\ 32 & 64 \end{bmatrix}$$

(ii.) Find the eigenvalues (λ_i) and eigenvectors $(\mathbf{v}^{(i)})$ of that product. The characteristic equation of the product is

$$\begin{aligned} \det\left(\mathbf{A}^H\,\mathbf{A} - \lambda\mathbf{I}\right) = \det &\begin{bmatrix} (16-\lambda) & 32 \\ 32 & (64-\lambda) \end{bmatrix} \\ &= \lambda^2 - 80\lambda = 0 \\ &\quad \implies \quad \lambda = 80\ ,\ 0 \end{aligned}$$

Then, for $\lambda_1 = 80$, $\left(\mathbf{A}^H\,\mathbf{A} - \lambda_1\mathbf{I}\right)\mathbf{v}^{(1)} = \mathbf{0}$ leads to the augmented matrix:

$$\left[\begin{array}{cc|c} -64 & 32 & 0 \\ 32 & -16 & 0 \end{array}\right] \quad \implies \quad \mathbf{v}^{(1)} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$$

where we have again normalized the eigenvector $\mathbf{v}^{(1)}$.

Similarly, for Then, for $\lambda_2 = 0$, $\left(\mathbf{A}^H\,\mathbf{A} - \lambda_2\mathbf{I}\right)\mathbf{v}^{(2)} = \mathbf{0}$ leads to the augmented matrix:

$$\left[\begin{array}{cc|c} 16 & 32 & 0 \\ 32 & 64 & 0 \end{array}\right] \quad \implies \quad \mathbf{v}^{(2)} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{bmatrix}$$

solution:

However, for the reduced SVD, we only need the eigenvectors for the nonzero eigenvalues. So, in this case, we can discard $\mathbf{v}^{(2)}$. Therefore, for the reduced SVD,

$$\hat{\mathbf{V}} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$$

where we utilize the "hat" to denote a reduced form.

(iii.) For each $\lambda_i \neq 0$, define $\sigma_i = \sqrt{\lambda_i}$, where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > 0$. So

$$\lambda_1 = 80 \quad \Longrightarrow \quad \sigma_1 = 4\sqrt{5}$$

(Since λ_2 is zero, we don't use it!). So

$$\hat{\boldsymbol{\Sigma}} = [\,4\sqrt{5}\,]$$

(iv.) For each $\sigma_i \neq 0$, define $\mathbf{u}^{(i)} = \frac{1}{\sigma_i}\mathbf{v}^{(i)}$, where $\mathbf{v}^{(i)}$ is the eigenvector associated with λ_i.

So, for $\sigma_1 = 4\sqrt{5}$, we have

$$\mathbf{u}^{(1)} = \frac{1}{\sigma_1}\,\mathbf{A}\,\mathbf{v}^{(1)} \quad \Longrightarrow \quad \mathbf{u}^{(i)} = \frac{1}{4\sqrt{5}} \begin{bmatrix} 2 & 4 \\ 2 & 4 \\ 2 & 4 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

In this case, because $\sigma_2 = 0$, we do not use

$$\mathbf{u}^{(2)} = \frac{1}{\sigma_2}\,\mathbf{A}\,\mathbf{v}^{(2)}$$

to compute $\mathbf{u}^{(2)}$, since the result would be undefined. (And, in general, we do not need $\mathbf{u}^{(i)}$ for the **reduced** SVD when $\sigma i = 0$.)

solution:

Therefore, in this case, we have

$$\hat{\mathbf{U}} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

Direct computation will verify that

$$\hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^H = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \left[4\sqrt{5} \right] \left[\begin{array}{cc} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{array} \right] = \begin{bmatrix} 2 & 4 \\ 2 & 4 \\ 2 & 4 \\ 2 & 4 \end{bmatrix} = \mathbf{A}$$

Furthermore, in this case, the full SVD would be

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4\sqrt{5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix}$$

17. Repeat the calculation of the SVD using the eigenvalue and eigenvector approach from problem 15 for the matrix:

$$\mathbf{A} = \begin{bmatrix} 16 & -2 \\ 13 & 14 \end{bmatrix}$$

solution:

First form the matrix:

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 16 & 13 \\ -2 & 14 \end{bmatrix} \begin{bmatrix} 16 & -2 \\ 13 & 14 \end{bmatrix} = \begin{bmatrix} 425 & 150 \\ 150 & 200 \end{bmatrix}$$

and find the characteristic polynomial:

$$P_3(\lambda) = \det \begin{vmatrix} 425-\lambda & 150 \\ 150 & 200-\lambda \end{vmatrix} = \lambda^2 - 625\lambda + 62500 = (\lambda - 125)(\lambda - 500)$$

$$\Longrightarrow \quad \lambda = 500, 125 \quad \Longrightarrow \quad \sigma = 10\sqrt{5}, 5\sqrt{5}$$

Find the associated eigenvectors. For $\lambda_1 = 500$:

$$\left(\mathbf{A}^T\mathbf{A} - \lambda_1 \mathbf{I}\right)\mathbf{v}^{(1)} = \mathbf{0} \quad \Longrightarrow \quad \begin{bmatrix} -75 & 150 & 0 \\ 150 & -300 & 0 \end{bmatrix} \quad \Longrightarrow \quad \hat{\mathbf{v}}^{(1)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

which normalizes to:

$$\mathbf{v}^{(1)} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

Similarly, for $\lambda_2 = 125$,

$$\left(\mathbf{A}^T\mathbf{A} - \lambda_2 \mathbf{I}\right)\mathbf{v}^{(2)} = \mathbf{0} \quad \Longrightarrow \quad \begin{bmatrix} 300 & 150 & 0 \\ 150 & 75 & 0 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{v}^{(2)} = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$$

(after normalization). And so

$$\mathbf{V} = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

(Observe the columns of $\mathbf{V}$ are obviously orthonormal!)

solution:

Next, construct the elements of $\mathbf{U}$ corresponding to the non-zero singular values of $\mathbf{A}$, i.e. to σ_1 and σ_2.
For $\sigma_1 = 10\sqrt{5}$,

$$\mathbf{u}^{(1)} = \frac{1}{\sigma_1}\mathbf{A}\mathbf{v}^{(1)} = \frac{1}{10\sqrt{5}}\begin{bmatrix} 16 & -2 \\ 13 & 14 \end{bmatrix}\begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$$

(Note that $\mathbf{u}^{(1)}$ is already normalized. This should have been expected!)

Then, for $\sigma_2 = 5\sqrt{5}$,

$$\mathbf{u}^{(2)} = \frac{1}{\sigma_2}\mathbf{A}\mathbf{v}^{(2)} = \frac{1}{5\sqrt{5}}\begin{bmatrix} 16 & -2 \\ 13 & 14 \end{bmatrix}\begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

(Note that, as also expected, $\mathbf{u}^{(2)}$ is already normalized.

At this point, $\mathbf{u}^{(1)}$ and $\mathbf{u}^{(2)}$ form a basis for $\Re^2$, so we're basically done, i.e.

$$\mathbf{U} = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

The corresponding matrix of singular values here is

$$\mathbf{\Sigma} = \begin{bmatrix} 10\sqrt{5} & 0 \\ 0 & 5\sqrt{5} \end{bmatrix}$$

Direct computation (MATLAB) will verify that

$$\mathbf{A} = \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^T$$

or

$$\begin{bmatrix} 16 & -2 \\ 13 & 14 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}\begin{bmatrix} 10\sqrt{5} & 0 \\ 0 & 5\sqrt{5} \end{bmatrix}\begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}^T$$

18. Repeat the calculation of the (reduced) SVD using the eigenvalue and eigenvector approach from problem 15 for the matrix:

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$$

solution:

First form the matrix:

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix}$$

and find the characteristic polynomial:

$$\begin{aligned} P_3(\lambda) = \det \begin{vmatrix} 10-\lambda & 6 \\ 6 & 10-\lambda \end{vmatrix} &= \lambda^2 - 20\lambda + 64 \\ &= (\lambda-4)(\lambda-16) \quad \Longrightarrow \quad \lambda = 16, 4 \quad \Longrightarrow \quad \sigma = 4, 2 \end{aligned}$$

Find the associated eigenvectors. For $\lambda_1 = 16$:

$$\left(\mathbf{A}^T\mathbf{A} - \lambda_1 \mathbf{I}\right)\mathbf{v}^{(1)} = \mathbf{0} \quad \Longrightarrow \quad \begin{bmatrix} -6 & 6 & 0 \\ 6 & -6 & 0 \end{bmatrix} \quad \Longrightarrow \quad \hat{\mathbf{v}}^{(1)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

which normalizes to:

$$\mathbf{v}^{(1)} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Similarly, for $\lambda_2 = 4$,

$$\left(\mathbf{A}^T\mathbf{A} - \lambda_2 \mathbf{I}\right)\mathbf{v}^{(2)} = \mathbf{0} \quad \Longrightarrow \quad \begin{bmatrix} 6 & 6 & 0 \\ 6 & 6 & 0 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{v}^{(2)} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

(after normalization). And so

$$\mathbf{V} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

solution:

Next, construct the elements of $\mathbf{U}$ corresponding to the non-zero singular values of $\mathbf{A}$, i.e. to σ_1 and σ_2.
For $\sigma_1 = 4$,

$$\mathbf{u}^{(1)} = \frac{1}{\sigma_1}\mathbf{A}\mathbf{v}^{(1)} = \frac{1}{4}\begin{bmatrix} 3 & 1 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

(Note that $\mathbf{u}^{(1)}$ is already normalized. This should have been expected!)

Then, for $\sigma_2 = 2$,

$$\mathbf{u}^{(2)} = \frac{1}{\sigma_2}\mathbf{A}\mathbf{v}^{(2)} = \frac{1}{2}\begin{bmatrix} 3 & 1 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

(Note that, as also expected, $\mathbf{u}^{(2)}$ is already normalized.

At this point, we still need on more orthonormal vector, $\mathbf{u}^{(3)}$, in order for the $\mathbf{u}^{(i)}$ to form a basis for $\mathbb{R}^3$. There is only one choice possible here, which we may find either by Gram-Schmidt, or $\mathbf{Q}\,\mathbf{R}$, or by inspection. That choice is:

$$\mathbf{u}^{(3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and therefore

$$\mathbf{U} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The corresponding matrix of singular values here is

$$\mathbf{\Sigma} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$$

solution:

Direct computation (MATLAB) will verify that

$$\mathbf{A} = \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^T$$

or

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T$$

19. The equation $\mathbf{A}\hat{\mathbf{V}} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}$ always implies that $\mathbf{A}\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}\,\hat{\mathbf{V}}^H$. However, if $\hat{\mathbf{V}}$ is "only" a **nonsquare** matrix with orthonormal columns, the latter does not immediately allow us to conclude the reduced SVD, i.e. that $\mathbf{A} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}\,\hat{\mathbf{V}}^H$, since, it it is quite likely that $\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H \neq \mathbf{I}$. However, if we can append enough additional orthonormal columns to $\hat{\mathbf{V}}$ to create a square matrix, **and if** all of those additional columns lie in Null($\mathbf{A}$), then we can still derive the reduced SVD. To see why this is true, assume $\mathbf{V} = \left[\begin{array}{cc} \hat{\mathbf{V}} & \hat{\hat{\mathbf{V}}} \end{array}\right]$, where all the columns of $\hat{\hat{\mathbf{V}}}$ lie in Null($\mathbf{A}$). Show that, in this case

$$\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \mathbf{I} - \hat{\hat{\mathbf{V}}}\,\hat{\hat{\mathbf{V}}}^H$$

and therefore $\mathbf{A}\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \mathbf{A}$.

solution:

Expressed in terms of block matrices, observe that

$$\mathbf{V} = \left[\begin{array}{cc} \hat{\mathbf{V}} & \hat{\hat{\mathbf{V}}} \end{array}\right] \quad \Longrightarrow \quad \mathbf{V}\,\mathbf{V}^H = \left[\begin{array}{cc} \hat{\mathbf{V}} & \hat{\hat{\mathbf{V}}} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{V}} \\ \hat{\hat{\mathbf{V}}} \end{array}\right]$$

But, because, with the added columns, $\mathbf{V}$ is now square with orthonormal columns, it is unitary, and $\mathbf{V}^H = \mathbf{V}^{-1}$. Therefore,

$$\mathbf{I} = \mathbf{V}\,\mathbf{V}^H = \left[\begin{array}{cc} \hat{\mathbf{V}} & \hat{\hat{\mathbf{V}}} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{V}} \\ \hat{\hat{\mathbf{V}}} \end{array}\right] = \hat{\mathbf{V}}\,\hat{\mathbf{V}}^H + \hat{\hat{\mathbf{V}}}\,\hat{\hat{\mathbf{V}}}^H$$

or, equivalently $\quad \hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \mathbf{I} - \hat{\hat{\mathbf{V}}}\,\hat{\hat{\mathbf{V}}}^H$

But then

$$\mathbf{A}\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \mathbf{A}\left(\mathbf{I} - \hat{\hat{\mathbf{V}}}\,\hat{\hat{\mathbf{V}}}^H\right) = \mathbf{A} - \mathbf{A}\hat{\hat{\mathbf{V}}}\,\hat{\hat{\mathbf{V}}}^H$$

However now, because every column of $\hat{\hat{\mathbf{V}}}$ lies in Null($\mathbf{A}$), then $\mathbf{A}\hat{\hat{\mathbf{V}}} = \mathbf{0}$, which immediately implies that the second term on the right in the expression immediately above must also vanish. Therefore,

$$\mathbf{A}\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \mathbf{A}$$